

Aberystwyth University

Optimization of Mixed Numerology Profiles for 5G Wireless Communication Scenarios

Correia, Noélia; Al-Tam, Farooq; Rodriguez, Jonathan

Published in:
Sensors

DOI:
[10.3390/s21041494](https://doi.org/10.3390/s21041494)

Publication date:
2021

Citation for published version (APA):

Correia, N., Al-Tam, F., & Rodriguez, J. (2021). Optimization of Mixed Numerology Profiles for 5G Wireless Communication Scenarios. *Sensors*, 21(4), 1-22. [1494]. <https://doi.org/10.3390/s21041494>

Document License CC BY

General rights

Copyright and moral rights for the publications made accessible in the Aberystwyth Research Portal (the Institutional Repository) are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Aberystwyth Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Aberystwyth Research Portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

tel: +44 1970 62 2400
email: is@aber.ac.uk

Article

Optimization of Mixed Numerology Profiles for 5G Wireless Communication Scenarios [†]

Noélia Correia ^{1,*} , Faroq Al-Tam ²  and Jonathan Rodriguez ³ 

¹ Center for Electronic, Optoelectronic and Telecommunications (CEOT), University of Algarve, 8005-139 Faro, Portugal

² Department of Computer Science, Aberystwyth University, Aberystwyth SY23 3FL, UK; faa14@aber.ac.uk

³ Instituto de Telecomunicações, Universidade de Aveiro, 3810-193 Aveiro, Portugal; jonathan@av.it.pt

* Correspondence: ncorreia@ualg.pt

[†] This paper is an extended version of the paper published in: Correia, N.; Al-Tam, F.; Rodriguez, J. Adaptive Spectrum Allocation for 5G Wireless Communication Scenarios. IEEE 25th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), Pisa, Italy, 14–16 September 2020, pp. 1–6, doi:10.1109/CAMAD50429.2020.9209267.

Abstract: The management of 5G resources is a demanding task, requiring proper planning of operating numerology indexes and spectrum allocation according to current traffic needs. In addition, any reconfigurations to adapt to the current traffic pattern should be minimized to reduce signaling overhead. In this article, the pre-planning of numerology profiles is proposed to address this problem, and a mathematical optimization model for their planning is developed. The idea is to explore requirements and impairments usually present in a given wireless communication scenario to build numerology profiles and then adopt one of the profiles according to the current users/traffic pattern. The model allows the optimization of mixed numerologies in future 5G systems under any wireless communication scenario, with specific service requirements and impairments, and under any traffic scenario. Results show that, depending on the granularity of the profiles, the proposed optimization model is able to provide satisfaction levels of 60–100%, whereas a non-optimized approach provides 40–65%, while minimizing the total number of numerology indexes in operation.

Keywords: 5th-Generation mobile networks; radio resource management; mixed numerology



Citation: Correia, N.; Al-Tam, F.; Rodriguez, J. Optimization of Mixed Numerology Profiles for 5G Wireless Communication Scenarios. *Sensors* **2021**, *21*, 1494. <https://doi.org/10.3390/s21041494>

Academic Editor: Gwanggil Jeon

Received: 23 December 2020

Accepted: 17 February 2021

Published: 21 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Boosting the data rate has been the main objective of wireless communication systems over this last decade. Long-Term Evolution (LTE) aimed at improving user data rate experience, and all the enhancements were centered on this main requirement, while the needs of services like virtual reality and the Internet of Things (IoT) were not properly taken into account. Currently, we are experiencing the 5th-Generation (5G) mobile era, where the objective is to allow the coexistence of different services and heterogeneous users on the same networking platform, besides looking for improvements of the data rate.

5G is a multi-use-case environment designed to reach almost all life aspects, from industry and health care to online games and vehicle-to-everything (V2X) communications. The current 5G is, however, just the beginning of more advanced communication systems. The service stack will continue to grow, and new applications are expected to be developed. It is no surprise that new requirements will also emerge in the near future, related to energy in particular, and more advanced designs will undoubtedly be needed [1].

The current 5G specifications describe it as a multi-purpose communication ecosystem. That is, it will be an umbrella for several services where each service is associated with a bundle of requirements (see [2] for further details). In order to support them, 5G is equipped with new features. Among them, flexibility is the most important one [3]. Supporting multiple use-cases requires a flexible design not only on radio access methods

and deployment options, but also on the frame structure and operations it supports. The corner stone to achieve such flexibility is the new design of the 5G new radio (NR).

The NR is built on four main pillars: (i) new spectrum; (ii) massive MIMO and beam-forming; (iii) multi-connectivity; (iv) network flexibility and virtualization (numerology, slicing, NFV/SDN). The flexibility of NR spreads all over the system planes. Not only on the frame structure, but also on the operation and protocol stack logic. Furthermore, NR is user-centric, and bandwidth parts (BWPs) can be used to fit user equipment (UE) temporal requirements with dynamic transmission time interval (TTI) lengths and dynamic time division duplex (DTDD) [4]. In short, the 5G NR is designed with components that are flexible, ultra-lean, and forward-compatible [5].

The NR frame structure is the pivotal element for the flexible support of heterogeneous services, while allowing adaptation to different user channel conditions [6]. Flexibility here comprises different orthogonal frequency division multiplexing (OFDM)-based waveforms, as well as a mixed numerology [7]. The last term refers to different multicarrier modulation parameters with an impact on subcarrier spacing, cyclic prefix (CP) duration, and slot duration, as shown in Table 1. The downlink and uplink transmissions are organized into 10ms frames, each having ten subframes of 1ms. In 5G, the 1ms subframe is then divided into one or more slots, depending on the numerology index in use. As shown in Figure 1, the shortening of the slot duration is related to the shortening of the duration of the symbols.

Table 1. Numerology structure in 5G.

Index	Δf (kHz)	T_{CP} (μs)	T_{slot} (ms)	Efficiency
0	15	4.7	1	93.4%
1	30	2.3	0.5	93.4%
2	60	1.2/4.2	0.25	93.4/80.0%
3	120	0.6	0.125	93.4%
4	240	0.3	0.0625	93.4%

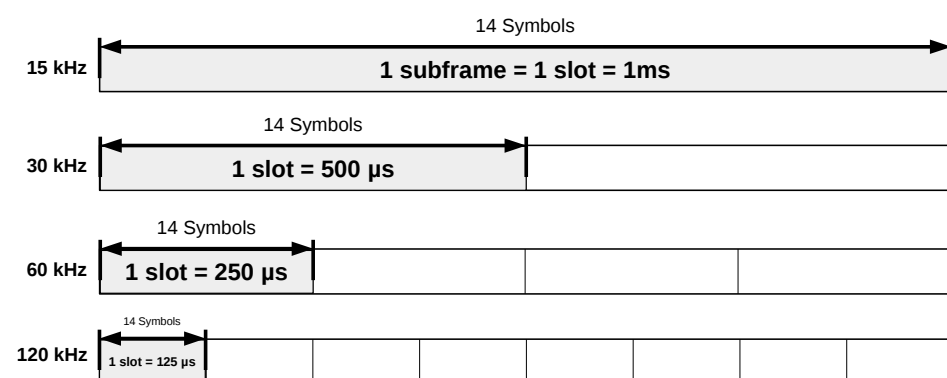


Figure 1. 5G new radio (NR) frame structure.

On the one hand, this flexible design enables an efficient delivery of different qualities like low latency, guaranteed bit rate (GBR), reliability, and more [8]. On the other hand, 5G resource management becomes more challenging. Mixed numerologies are prone to: (i) inter-numerology interference (INI); (ii) low spectral efficiency; (iii) signaling overhead; (iv) scheduling complexity. Therefore, the number of coexisting numerology indexes should be minimized.

A mixed numerology system where different BWPs co-exist demands special planning. Such a planning problem involves deciding on the numerology mix and BWPs that better explore network capacity given a set of users/services that may change in number, requirements, and impairments. Such planning should also be feasible for environments

having a dynamic traffic pattern. Here in this article, the pre-planning of the numerology profiles is proposed to address this issue. The mathematical optimization models created to design such mixed numerology profiles allow for the optimization of mixed numerologies in future 5G systems, under any wireless communication scenario and traffic pattern. More clearly, the contributions of this article are the following:

- A framework for the outline of mixed numerology profiles is proposed. These profiles are planned according to the k most antagonistic user/service requirements. The proposed framework allows a fast transition between profiles, performed in the case of traffic pattern changes.
- A two-step approach is proposed for the planning of such numerology profiles, each having its own mixed numerology and BWPs. Mathematical optimization models are presented to solve both steps.

As technology evolves in 5G, operators will be able to deliver even more advanced and value-added services. Therefore, network planning and deployments must be done in such a way that they match the ambition of the services. This requires understanding users/services and taking into account their requirements in order to outline future mixed numerology profiles. The proposed framework allows this goal to be reached.

The remainder of this article is organized as follows. Section 2 discusses related work, and then, Section 3 clarifies several wireless communication and 5G related definitions, which are required to understand the following sections and addressed problem. In Section 4, the motivation for the optimization of numerology profiles is presented, and a mathematical formulation of such an optimization problem is created, allowing it to be solved. Section 5 presents the simulation setup and analysis of the results, and Section 6 presents a final discussion on the results and conclusions, together with future work.

2. Related Work

The multiplexing of 5G services under a mixed numerology is now a hot topic. The mixture of OFDM-based numerology indexes was analyzed in [6], where it was shown to be beneficial for the support of different services with different latencies. However, due to energy leaks, the INI problem may appear. The authors demonstrated a time-domain window filtering method to reduce such an effect. The authors in [9] studied the multi-service support under different subcarrier spacing and highlighted the effect of the subcarrier spacing difference and the guard band size on the INI problem. They proposed a sub-band filtered transmission scheme and cancellation and equalization algorithms to reduce the effect of the INI problem. In [10], the INI problem was also analyzed and shown to be even more severe on the edge sub-carriers of neighboring BWPs. Similar to [9], the authors also showed that the effect of the INI is proportional to the difference in numerology indexes. In order to avoid the INI, an adaptive numerology selection approach was developed in [11] based on the delay requirements of the service. This approach considers only the delay requirements to reduce the average scheduling latency.

Finding the minimum mixture of numerology indexes was studied in [12], where a metric was developed to measure user satisfaction regarding the numerology assigned to deliver its services. Different scenarios were simulated and different numerology index sets considered. In order to select the desired number of numerologies, the authors proposed a greedy algorithm, which uses the metric to obtain a trade-off between scenario and user requirements. In [13], service multiplexing using a predefined mixture of numerology indexes was modeled as an integer programming problem and shown to be NP-hard by constructing a polynomial-time reduction from the partition problem. For this reason, it was solved using a Lagrangian relaxation. However, this work did not include the INI problem in the model. Similar studies can be found in [14,15]. In [14], a multi-user OFDM system was modeled using an integer program, and a relaxed version was developed to reduce the computational complexity. In [15], an integer program was also developed to model the scheduling of a multi-user system while meeting users' service requirements. This work also proved that the multi-user resource allocation under different service

requirements is an NP-hard problem. Moreover, the authors developed two different algorithms: a resource portioning algorithm to decompose the allocation problem into a set of parallel small scale problems and an iterative greedy algorithm based on the resource assignment weight. This work, however, allocated physical resource blocks (PRBs) to users under the assumption that numerologies are preselected by each user. The energy efficiency optimization problem was studied in [16]. Furthermore, a joint beamforming and power allocation scheme was proposed, which takes into consideration the intra- and inter-cell interference, but not the INI problem. In addition, resource scheduling of different services was studied in [17], where different models were developed to join enhanced mobile broadband (eMBB) and ultra reliable low latency communications (URLLC) resource allocation. In [18], the concern was to predict the outage probability in order to ensure efficient and stable service communications. New approaches using machine learning have also emerged, as in [19]. Table 2 summarizes the differences among these research efforts.

The previously mentioned studies considered either a predefined mixed numerology and addressed resource allocation and service multiplexing problems or considered planning a mixed numerology that is more suitable for specific services. In [20], we addressed the spectrum allocation problem. In this article, and contrary to such studies, the problem is to plan multiple numerological profiles that fit certain QoS requirements, which should be selected according to the presence or not of one or more of these QoS requirements over time. To our knowledge, this problem has not been addressed before.

Table 2. Mixed numerology related work.

Ref.	Year	Main Focus	Outcome	Pros and Cons
[6]	2016	Analyzing the mixture of numerologies to support different services.	Showed that multiple numerologies are beneficial for supporting different services and proposed a time-domain window filtering (W-OFDM) method to reduce the effect of the inter-numerology interference (INI).	W-OFDM suffers from imperfect power amplifier (PA) nonlinearity.
[9]	2017	Multi-service support under different subcarrier spacing.	Highlighted the effect of the subcarrier spacing difference and the guard band size on the severity of the INI problem and proposed a sub-band filtered transmission scheme using cancellation and equalization algorithms to reduce the effect of the INI problem.	Accurate channel state and noise variance feedback are required for proper cancellation [21].
[10]	2018	The INI problem effect across the sub-carriers.	Showed that the INI is more severe on the edge sub-carriers of neighboring bandwidth parts (BWPs).	Only simulation results under systems with adjacent bands were analyzed.
[16]	2018	Energy efficiency optimization.	A scheme that joins beamforming and power allocation.	Considered the intra- and inter-cell interference, but not the INI problem.
[12]	2018	The minimum mixture of numerologies for a scenario.	A metric is to measure user satisfaction, and a greedy algorithm to obtain a trade-off between scenario and user requirements.	A simple metric and algorithm, but not optimal.
[13]	2018	Modeling the service multiplexing using a predefined mixture of numerology indexes.	An integer program, and a Lagrangian relaxation solution.	Did not consider the INI in the model.
[11]	2019	Adaptive numerology selection.	An adaptive selection approach able to reduce the average scheduling latency.	Considered only the delay requirements.
[14]	2019	Modeling the multi-user OFDM system under a mixture of numerologies.	An integer program and a relaxed solution.	Assumed a fixed mixture of numerologies.

Table 2. Cont.

Ref.	Year	Main Focus	Outcome	Pros and Cons
[20]	2020	Resource allocation under different services and a mixture of numerologies.	An optimization model (a linear program).	Can solve the problem in polynomial time, but did not include profile planning.
[15]	2020	Modeled the scheduling of a multi-user system under different service requirements.	An integer program, a resource partitioning algorithm, and an iterative greedy algorithm.	Allocated physical resource blocks (PRBs) to users under the assumption that numerologies are preselected by each user.
[17]	2020	Modeled the joint enhanced mobile broadband (eMBB) and ultra reliable low latency communications (URLLC) scheduling and analyze the eMBB rate loss associated with URLLC superposition.	Different loss function models: linear, convex, and threshold; and a solution for each model.	The accuracy of the linear model was not high in practice [22].
[19]	2020	Resource allocation under different numerologies.	A deep reinforcement learning solution for resource allocation.	Did not consider a mixture of numerologies.

3. Required Definitions

For readability, the notation used throughout the article is summarized in Table 3.

Table 3. Notation. CQI, channel quality information; DRB, data radio bearer.

Symbol	Meaning
B	Frequency Band Covering the Service Area.
W	Maximum transmission bandwidth covering the service area.
$\mathcal{G}(B, W)$	Set of allowed numerology indexes for frequency band B under bandwidth W ; g denotes an element of this set.
\mathcal{P}_u	Set of BWPs for utilization by UE u ; p denotes an element of this set.
$\text{perc}(p)$	Bandwidth percentage assigned with BWP p .
C_u	Transfer data limitation for UE u , per 66.67us(symbol duration in Numerology Index 0).
m	Modulation data transfer factor.
\mathcal{I}	Set of measurements/feedback provided via CQI or other similar systems.
\mathcal{S}	Set of services.
\mathcal{Q}	Set of QoS requirements; q denotes an element of this set.
\mathcal{D}	Set of DRBs in a wireless communication scenario.
\mathcal{D}_u	Set of DRBs for UE u ; d denotes an element of this set.
\mathbf{r}^d	Vector of QoS requirement values for DRB d ; $r^d(q)$ denotes the value of requirement q .
\mathbf{e}_g^d	Vector with the relative effect of numerology index g on QoS requirements, for DRB d ; $e_g^d(q)$ denotes the effect on requirement q , $0 \leq e_g^d(q) \leq 1$.
$\Delta(d, g)$	The most critical QoS requirement, for DRB d , under numerology index g .
$\Phi(d, g, q)$	Impact of numerology g on QoS requirement q , for DRB d .
\mathcal{D}_k	Set of all size- k subsets of \mathcal{D} ; \mathcal{D}_k^i denotes one of the subsets, and $\mathcal{D}_k^i[l]$ denotes a DBR in subset \mathcal{D}_k^i .
$\bar{\mathcal{D}}$	Most antagonistic DRBs.
k	Maximum number of competing DRBs to be considered.
\mathcal{P}	Set of numerology profiles, of size 2^k ; p denotes one of the profiles.
P	Number of numerology profiles.
\mathcal{G}_p	Indexes used in numerology profile p .

For now, 5G systems will use a single waveform (CP-OFDM), and two large frequency ranges (FRs) are specified by 3GPP: sub-6 GHz (FR1) and millimeter wave (FR2). A subcarrier spacing of 15 and 30 kHz can be used in sub-6 GHz; a subcarrier spacing of 120 kHz

can be used in the millimeter wave range; and a subcarrier spacing of 60 kHz can be used in both. A set of bands has been defined for each FR, by 3GPP, together with the available subcarrier spacings and supported UE channel bandwidths (maximum transmission bandwidth + guard bands). For example, a 5 MHz UE channel bandwidth is only supported in a 15 kHz subcarrier spacing [23].

Definition 1 (Cell bandwidth). *Let us assume a wireless communication scenario in which the service area is covered by a specific frequency band B , under a certain bandwidth denoted by W . Guard band penalties are assumed to be incorporated in W , and the set of allowed numerology indexes is denoted by $\mathcal{G}(B, W)$.*

The cell bandwidth is expected to be large, but the reception/transmission bandwidth of a UE is not necessarily the same as that of the cell bandwidth. That is, the reception/transmission bandwidth of a UE will be a subset of the total cell bandwidth and may decrease during low activity to save power. A UE can have at most four BWPs configured for downlink (similar for uplink; supplementary uplink is possible), and for now, just a single BWP will be active at a time (in Release 15). One of the BWPs will be similar for all users in the cell and used for initial access to the network.

Definition 2 (BWP). *A bandwidth part is the frequency spectrum, within the carrier's bandwidth, over which a device is currently operating. A BWP is a group of contiguous physical resource blocks (PRBs) where one PRB occupies 12 consecutive subcarriers (frequency domain), and it can be used in either direction (uplink or downlink). A BWP is associated with a numerology index, and only a single BWP can be active at any time. The set of BWPs for utilization by UE u is denoted by \mathcal{P}_u , and if $\text{perc}(p)$ is the bandwidth percentage assigned to $p \in \mathcal{P}_u$, $0 \leq \text{perc}(p) \leq 1$, then $\lfloor \frac{\text{perc}(p) \times W}{0.18 \times 2^g} \rfloor$ will be the number of its PRBs ($0.18 \text{ MHz} = 15 \text{ kHz} \times 12$) if numerology index $g \in \mathcal{G}(B, W)$ is used. A switch to short BWP allows energy saving, while BWPs at different numerology indexes allow for different services.*

The control mechanism responsible for exchanging BWP information is the radio resource control (RRC) protocol. This protocol is understood by both the user NR and by the network gNB. The RRC can perform BWP reconfiguration or use downlink control information (DCI) messages for BWP switching. These consume a certain time, and for this reason, their use should be minimized. Since the RRC reconfiguration is the most demanding one, the assigned BWPs should fit not only current user/service needs (while being efficient regarding the use of physical resources), but also future needs so that RRC reconfigurations are avoided as much as possible.

Definition 3 (NR data rate). *A user NR device is assumed to have a limitation on the speed at which it can transfer data, denoted by C_u . It is assumed that such a limitation is per $66.67 \mu\text{s}$ (symbol duration in Numerology Index 0). Such a limitation influences the user BWP sizes, numerology, and modulation scheme. That is, for a given BWP of user u , $\#PRBs \times 12 \times m \times (g + 1) \leq C_u$, for numerology g and modulation m .*

Moreover, NR devices also have a limitation on the speed at which they can transfer data, and impairments (e.g., Doppler spread, frequency offset) may exist.

Definition 4 (Impairments). *Measurements/feedback are provided via channel quality information (CQI), or other similar systems. The overall set of feedback elements is denoted by \mathcal{I} , and it is assumed that specific wireless communication scenarios (e.g., mobile), served by a specific frequency band, usually lead to a set of channel and UE impairments. Such information should be used to improve the QoS.*

ITU-R started by defining three service types: enhanced mobile broadband (eMBB), ultra reliable low latency communications (URLLC), and massive machine type communications (mMTC) [24]. Regarding these service types, it can be stated that:

- eMBB: The focus is on supporting the ever-increasing end user data rate and system capacity.
- URLLC: The focus is reliability and security required by mission-critical applications. High subcarrier spacing and mini-slots are the key enablers for this use case.
- mMTC: The focus is energy efficiency and massive connectivity.

This classification has as a basis the technical viewpoint of network operators and service providers. More recently, an end user experience perspective was proposed to classify 5G services [25,26]. The authors proposed a classification based on five features: immersiveness, intelligence, omnipresence, autonomy, and publicness. Naturally, other kinds of services may arise, allowing the possibility to better serve existing and future applications. As the number of classes increases, so does the complexity of systems because services will have different kinds of requirements.

The service requirements are specified through 5G QoS class index (5QI) values [27,28], and these are expected to differ among wireless communication scenarios like indoor hotspot, dense urban, rural, urban macro, high speed, etc. [29]. That is, wireless communication scenarios are expected to affect the service requirements, which is basically related to the impairments.

Definition 5 (Services). *The main types of uses that 5G is expected to enable (overall set of services under consideration) are denoted by S . The overall set of possible QoS requirements is denoted by \mathcal{Q} , and it is assumed that requirement values change according to the wireless communication scenario.*

Regarding the duration of TTIs, these will change according to the used numerology index because $\text{TTI} = \text{the number of symbols in time} \times \text{symbol length}$. For services requiring lower latencies, a low number of symbols per TTI or a short symbol length can be used to obtain a shorter TTI. For higher spectral efficiency, a longer TTI allows for higher spectral efficiency (less % of downlink control channel overhead, used to carry scheduling decisions every TTI). Therefore, there is a tradeoff between system spectral efficiency and minimal latency.

In practice, the subcarrier spacing, TTI, and number of subcarriers are directly related. For a low number of subcarriers, large subcarrier spacing is used, and a lower TTI (due to short symbol duration) is obtained, leading to low latency (note, however, that TTI has to do also with the number of symbols in time). Users/services can be flexibly multiplexed over the available resources with different TTIs, which allows for the support of service-aware TTI multiplexing on the same frequency [30]. The TTI can be adjusted according to the required latency and scheduling frequency [31].

The radio medium access control (MAC) scheduler does the packet treatment separately for each data radio bearer (DRB) following a two-step mapping of end-to-end (E2E) session flows.

Definition 6 (Two-step mapping of E2E session flows). *The non-access stratum (NAS) filters the data packets in the UE, or 5G core network (CN), and associates the data packets with QoS flows. An E2E session can be associated with one or more QoS flows. The access stratum (AS) mapping in the UE (or 5G RAN) associates the QoS flows with the DRBs (see Figure 2). This mapping is based on 5QI in the transport header of the packets and on the corresponding QoS parameters that are signaled via the CN interface when a session is established. One or more QoS flows can be mapped to a DRB, and a UE can have a set of DRBs, denoted by $\mathcal{D}_u = \{d_u^1, \dots, d_u^{|\mathcal{D}_u|}\}$. A DRB d has a vector of QoS requirement values denoted by $\mathbf{r}^d = [r^d(q)]$, where $q \in \{1, \dots, Q\}$.*

Table 4 shows some traffic types and corresponding requirements. If the system reaches congestion, then priorities can always be used (see [28] for details). This two-

step approach allows differentiating application/service flows and the adaptation of the DRB requirements to guide the radio scheduler. That is, the QoE manager can adaptively monitor and adjust the mapping of QoS flows to DRBs (reflective QoS). By adjusting the mapping of QoS flows to a DRB, there is an adjustment of the latency budget, packet loss rate tolerance, and GBR with the DRB, and this can be used to guide the lower-layer scheduler (although this should be done rarely, and priority should be changed to better serve an application). DRB data can be mapped to one or more BWPs and, consequently, TTI sizes.

Table 4. Requirements for traffic types. 5QI, 5G QoS class index.

Application	Latency (ms)	Avg Data Rate (Mb/s)	Guaranteed Bit Rate (Mb/s)	Packet Error Rate	5QI Value
VoIP	100	0.16	0.112	10^{-2}	1
Video conference	150	1	0.8	10^{-3}	2
Real-time gaming	50	0.8	0.72	10^{-3}	3
Buffered streaming	300	3.33	-	10^{-6}	6
Live streaming	100	1	-	10^{-3}	7
Pre-AR/VR	10	2	-	10^{-6}	80

Applications can have different requirements, and achieving the highest data rates may not always be the main requirement. Power consumption can be a critical issue and must also be considered. For most devices, the maximum throughput scenario is the one leading to the highest energy efficiency because the energy consumed per transferred bit is minimum (there is a power baseline that if distributed by a large number of bits, then the power efficiency is higher). However, it cannot be taken for granted that the power consumption will linearly map to a data rate improvement, especially if the device uses the full bandwidth to transmit lower data rate traffic (high baseline). That is, if the device uses a large bandwidth to transmit low data rate traffic, then there is energy inefficiency. In summary:

- Uplink: Since power varies according to the distance, when transmit power is near its limit, then the only way to extend uplink coverage is to concentrate the same energy into fewer bits. If the data are not urgent, then the BWP should be small. For short distances, higher BWPs (fitting data rate) should be used for higher energy efficiency).
- Downlink: High throughputs can be provided by high bandwidth carriers and MIMO layers, but this requires high processing capacity to deal with data rates and high power at maximum throughput because the device has to actively monitor wideband control channel across a large bandwidth, even when no data are present. For this reason, the BWP should be large in the case of high throughput, but small in the case of low throughput.

The power baseline also changes with the channel/BWP/modulation, and this will change the required energy per bit. Switching between BWPs can be done to keep the baseline low, if the throughput reduces, but the switching overhead must also be considered. In summary, the BWP and the consequent power consumption must be considered based on the traffic profile: service type and their requirements.

Definition 7 (NR energy). *NR devices are assumed to have different hardware characteristics. In general, the higher the channel, BWP, and modulation of a channel, the higher the power baseline. This should follow service throughput needs for energy efficiency. Full spectrum and short TTI allow constrained devices to go to sleep mode. BWP planning should take the traffic profile into account.*

Numerology indexes are not associated with specific service classes because there will be different communication scenarios and UE characteristics. For this reason, seven numerology indexes have been defined, and the BWP/numerology index in use by a UE may change at

every TTI. This should be decided according to: (i) the requirements of services/DRBs in a communication scenario; (ii) NR features (bandwidth, data rate, and energy limitations); and (iii) impairments.

In general, few numerological indexes should be used when there is a need for greater spectral efficiency. This is because a mixture of numerology indexes requires guard bands to avoid the INI. In some cases, such spectral efficiency is more important than having multiple numerology indexes, which provides more flexibility, but at the expense of scheduling complexity and signaling overhead, leading to a waste of bandwidth resulting from the use of multiple guard bands.

4. Management of Radio Resources

4.1. Motivation

As previously described, the reception/transmission bandwidth of a UE is expected to be smaller than the cell bandwidth. Therefore, different numerology indexes and BWP sizes will co-exist. That is, at a given time, there is a numerology mix and combination of active BWPs at the network. The problem is, therefore, to decide on the numerology mix and BWPs that better explore network capacity given a set of users that may change in number, requirements, and impairments. Here, in this article, the pre-planning of numerology profiles is proposed to address this problem. More specifically, the idea is to explore requirements and impairments usually present in a given wireless communication scenario to build numerology profiles and then adopt one of the profiles according to the current users/traffic pattern.

4.2. Defining the Numerology Profile Problem

To better understand what is meant by numerology profiles, some definitions need to be introduced.

Definition 8 (Critical QoS requirement). Let $\mathbf{r}^d = [r^d(q)]$ denote the vector of QoS requirements for a DRB d , $\forall q \in \mathcal{Q}$, where \mathcal{Q} is the set of QoS requirements. Each element in \mathbf{r}^d is the relative need for QoS requirement q that results from QoS flows mapped to DRB d . Let $\mathbf{e}_g^d = [e_g^d(q)]$ denote a vector with the relative effect of numerology index g on q , $\forall q \in \mathcal{Q}$. Then, the most critical QoS requirement for DRB d , under numerology index g , is given by:

$$\Delta(d, g) = \arg \max_{q \in \{1, \dots, Q_d\}} \{\Phi(d, g, q)\}, \quad (1)$$

where $\Phi(d, g, q) = r^d(q) - r^d(q) \times e_g^d(q)$ measures the impact of numerology g on requirement q , for DRB d , assuming $0 \leq e_g^d(q) \leq 1$.

The most critical QoS requirement changes, therefore, according to the numerology index under utilization. The relative effect of numerology index g on QoS requirement q can be extracted from Table 5, and multiple requirements may have to be simultaneously considered. This is detailed when discussing the simulation setup in Section 5. A specific numerology can also have an opposite effect on different DRBs. That is, there will be antagonistic DRB critical requirements present at the network.

Definition 9 (Antagonistic DRBs). Let us assume a wireless communication scenario with a representative set of DRBs with QoS requirements, denoted by \mathcal{D} . Assuming \mathcal{D}_k is the set of all size- k subsets of \mathcal{D} , then the k most antagonistic DRB critical requirements result from:

$$\arg \max_{\mathcal{D}_k^i \in \mathcal{D}_k} \left\{ \sum_{l \in \{1 \dots k\}} \max_{g \in \mathcal{G}} \left\{ \sum_{m \neq l} |\Phi(d, g, q) - \Phi(d', g, q')| \right\} \right\} \quad (2)$$

where $d = \mathcal{D}_k^i[l]$, $d' = \mathcal{D}_k^i[m]$, $q = \Delta(d, g)$, and $q' = \Delta(d', g)$.

Both \mathbf{r}^d and \mathbf{e}_g^d values, used in $\Delta(\cdot)$, are normalized between zero and one. Having these definitions in mind, the numerology profile problem can be defined as follows.

Definition 10 (Numerology profile problem). *Given a wireless communication scenario, with a representative set of DRBs with QoS requirements, denoted by \mathcal{D} , then $P = 2^k$ numerology profiles must be planned according to the presence of one or more of the k most antagonistic DRB critical requirements (possible combination of present DRBs). Their presence, at any moment in time, will lead to the adoption of one of the profiles by the network. Then, for each profile, it is necessary to decide on: (i) the most suitable numerology indexes; (ii) the percentage of spectrum to be allocated to each numerology index. Such a decision should try to ensure an optimal use of resources while providing fairness among users/services.*

Thus, given a wireless communication scenario, the numerology profiles reflect the anticipation of possible competitions for the spectrum by the k most antagonistic DRBs. Such planning criteria are flexible, using all degrees of freedom to reach the best multiplexing, depending on the presence (or not) of the k most antagonistic DRB critical requirements. Note that a high k value leads to fine-grained profiles.

Table 5. Numerology design needs for service requirements. TTI, transmission time interval; GBR, guaranteed bit rate.

Requirement	Δf (kHz)	# of Subcarriers	TTI Duration	T_{CP} (μs)	Spectral Efficiency	Comment
High data rate					High	Less scheduling control info
GBR			Short			High scheduling frequency
Latency	Large		Short			Short symbol duration
Low packet error	Large			Long		Low inter-symbol interference
Energy efficiency		Low	Short			Low baseline and sleep mode increase

Regarding impairments, it is assumed that these are incorporated in \mathbf{e}_g^d because certain wireless communication scenarios are known to exhibit specific impairments. Note also that due to the previously mentioned two-step mapping of E2E session flows, the requirements of DRBs in each wireless communication scenario are not expected to change much. Therefore, such a kind of long-term planning can always be applied to any wireless communication scenario.

4.3. Solving the Numerology Profile Problem

According to Definition 10, the numerology profile problem can be solved using a two-step approach (Figure 2). First, the numerology indexes that better serve the k most antagonistic DRB critical requirements, in each profile, should be found. Then, spectrum must be allocated to such numerology indexes, taking resource optimization and fairness into account. These two steps are discussed next.

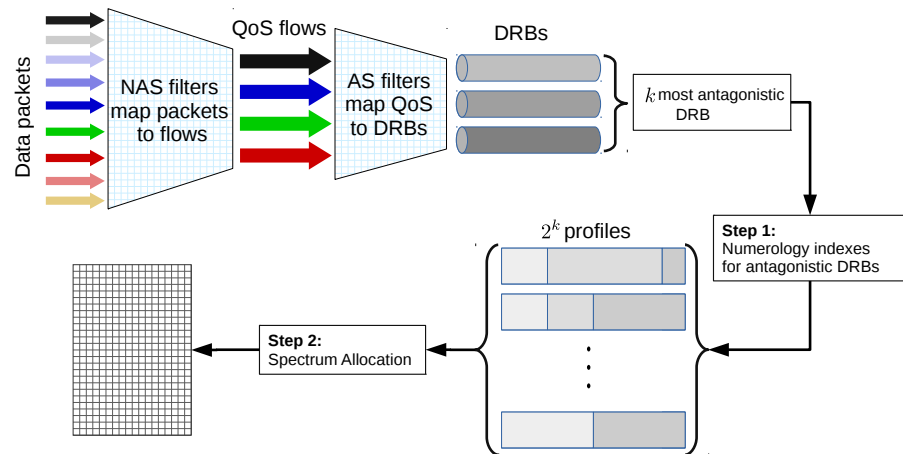


Figure 2. The two-step solution of the numerology selection problem. NAS, non-access stratum.

4.3.1. Most Antagonistic DRBs

Let us assume the following given information for a wireless communication scenario:

- \mathcal{G} Set of numerology indexes that can be used in the wireless communication scenario's frequency band, where g denotes a specific numerology index.
- \mathcal{D} Set of representative DRBs, where d denotes a specific DRB.
- k Maximum number of competing DRBs (granularity of profiles).
- \mathcal{Q} Set of QoS requirements.
- \mathbf{r}^d Vector with $|\mathcal{Q}|$ relative QoS needs of DRB d .
- \mathbf{e}_g^d Vector with the effectiveness of numerology g to accomplish the QoS needs of DRB d .

The variables are:

- λ^d One if DRB $d \in \mathcal{D}$ is considered one of the k most antagonistic DRB critical requirements; zero otherwise.
- $\Delta_{q,g}^d$ One if $q \in \mathcal{Q}$ is the most critical QoS requirements when DRB $d \in \mathcal{D}$ is under numerology index $g \in \mathcal{G}$; zero otherwise.
- $\gamma_{q,q',g}^{d,d'}$ One if critical requirements q and q' of antagonistic DRB (d, d') pair, under numerology g , is to be considered when evaluating the expression in Definition 10; zero otherwise.
- $\mu_{q,q',g}^{d,d'}$ Difference between critical requirements q and q' of antagonistic DRB (d, d') pair, under numerology g ($|\dots|$ in Equation (2)).

Given such notation, this step can be solved by the following mathematical optimization problem:

Objective function:

$$\text{maximize} \sum_{\{d,d' \in \mathcal{D}\}} \sum_{\{q \in \mathcal{Q}_d\}} \sum_{\{q' \in \mathcal{Q}_{d'}\}} \sum_{\{g \in \mathcal{G}\}} \mu_{q,q',g}^{d,d'} \quad (3)$$

A numerological index g will have different suitability degrees for each DRB QoS requirement. For fairness among users/services, the DRBs with the greatest difference in suitability (given a specific numerology index) should be the ones to take into account when choosing the numerology indexes in operation. This objective function searches for those opposite suitabilities, or most antagonistic DRBs.

Competing DRBs:

$$\sum_{\{d \in \mathcal{D}\}} \lambda^d = k \quad (4)$$

This constraint states that k DRBs, the most antagonistic, must be determined. This depends on the supported numerology index mixing level.

Critical QoS requirement:

$$\sum_{\{q \in \mathcal{Q}\}} \Delta_{q,g}^d = \lambda^d, \forall d \in \mathcal{D}, \forall g \in \mathcal{G} \quad (5)$$

Constraints (5) are used to ensure that a single QoS requirement is marked as the critical one, when DRB d is operating in numerology number g and is one of the most antagonistic DRBs. The right QoS requirement, to be considered as critical, is influenced by the objective function maximization and the next constraints.

Most competing DRBs:

$$\gamma_{q,q',g}^{d,d'} \leq \Delta_{q,g}^d, \forall d, d' \in \mathcal{D}, \forall q \in \mathcal{Q}_d, \forall q' \in \mathcal{Q}_{d'}, \forall g \in \mathcal{G} \quad (6)$$

$$\gamma_{q,q',g}^{d,d'} \leq \Delta_{q',g}^{d'}, \forall d, d' \in \mathcal{D}, \forall q \in \mathcal{Q}_d, \forall q' \in \mathcal{Q}_{d'}, \forall g \in \mathcal{G} \quad (7)$$

$$\mu_{q,q',g}^{d,d'} = \gamma_{q,q',g}^{d,d'} \times |[r^d(q) - r^d(q) \times c_g^d(q)] - [r^{d'}(q') - r^{d'}(q') \times c_g^{d'}(q')]|, \\ , \forall d, d' \in \mathcal{D}, \forall q \in \mathcal{Q}_d, \forall q' \in \mathcal{Q}_{d'}, \forall g \in \mathcal{G} \quad (8)$$

Constraints (6) and (7) find (d, d') pairs of antagonistic DRB critical requirements, those that must be considered when evaluating the expression in Definition 10. Their suitability difference, when using g , is stored in (8). Note that variables λ^d , $\Delta_{q,g}^d$, $\gamma_{q,q',g}^{d,d'}$, and $\mu_{q,q',g}^{d,d'}$ depend on each other, and their values are decided globally, as a whole, guided by the maximization goal as the objective function.

Non-negativity assignment to variables:

$$\lambda^d, \Delta_{q,g}^d, \gamma_{q,q',g}^{d,d'} \in \{0, 1\}; \mu_{q,q',g}^{d,d'} \geq 0. \quad (9)$$

4.3.2. Spectrum Allocation

After determining the most antagonistic DRBs, spectrum allocation must be performed while ensuring fairness among users/services. This requires introducing some assumptions based on the discussion in Section 4. These assumptions will serve as a basis for the optimization problem presented next.

Assumption 1 (Q1: Latency and GBR). *Services with latency and GBR limitations should have a short TTI, allowing a high scheduling frequency. Large BWPs (for less symbols in time) and/or high numerology indexes should be adopted.*

Assumption 2 (Q2: Data rate). *Services requiring high data rates should have a large channel bandwidth. A large TTI, for high spectral efficiency (less downlink scheduling control info), and a high modulation scheme should be used.*

Assumption 3 (Q3: Packet error rate). *For services requiring low packet loss, extended CP and low modulation schemes should be applied. The first acts as a guard band between successive symbols, overcoming inter-symbol interference (ISI) and enabling reliable operation.*

Now, let us assume the following known information:

$\bar{\mathcal{D}}$	Antagonistic DRBs determined in Step 1: $\bar{\mathcal{D}} = \{d \in \mathcal{D} : \lambda^d = 1\}$.
\mathbf{r}^d	Vector with relative QoS needs of DRB $d \in \bar{\mathcal{D}}$.
\mathbf{e}_g^d	Vector with the effectiveness of numerology g to accomplish the QoS needs of DRB d .
k	Maximum number of competing DRBs (granularity of profiles).
\mathcal{P}	Set of spectrum uses/profiles, of size 2^k , where p denotes one of the profiles.
a_i^d	One if DRB $d \in \bar{\mathcal{D}}$ is associated with position $i \in \{1 \dots k\}$; zero otherwise.
c_p^i	One if position $i \in \{1 \dots k\}$ is active in profile $p \in \mathcal{P}$; zero otherwise.
$\bar{\mathcal{G}}_p$	Numerology indexes for profile p given by $\bar{\mathcal{G}}_p = \{g : g = \arg \min_{g \in \mathcal{G}} \{\max_{q \in \mathcal{Q}_d} \{r^d(q) - r^d(q) \times e_g^d(q)\}\}, \forall d : a_i^d \times c_p^i = 1\}$.

Regarding $\bar{\mathcal{G}}_p$, this will store the best numerology indexes for each $d \in \bar{\mathcal{D}}$. More specifically, \max gives the worst impact of numerology index g , from all QoS requirements q of DRB d , and $\arg \min$ chooses the lowest value (numerology index with the least reduction of QoS). The variables are:

$\alpha_{p,g}^L$	Lowest PRB for numerology index $g \in \bar{\mathcal{G}}_p$ in profile $p \in \mathcal{P}$.
$\alpha_{p,g}^H$	Highest PRB for numerology index $g \in \bar{\mathcal{G}}_p$ in profile $p \in \mathcal{P}$.
Θ_p	Lower bound on the numerology index spectrum allocations in profile $p \in \mathcal{P}$.

Given such notation, the spectrum allocation can be modeled by the following mathematical optimization problem:

Objective function:

$$\text{maximize } \sum_{\{p \in \mathcal{P}\}} \Theta_p \quad (10)$$

In this second step, the objective is to maximize the lower bound on the numerology index spectrum allocation, for each profile, in order to ensure fairness among spectrum allocations.

Allocation of PRBs:

$$\alpha_{p,g}^L \leq \alpha_{p,g}^H, \forall p \in \mathcal{P}, \forall g \in \bar{\mathcal{G}}_p \quad (11)$$

$$\sum_{\{g \in \bar{\mathcal{G}}_p\}} (\alpha_{p,g}^H - \alpha_{p,g}^L) \times 2^g + [B \times (|\bar{\mathcal{G}}_p| - 1)] \leq W, \forall p \in \mathcal{P} \quad (12)$$

where B and W are, respectively, the numerology guard bands and the total bandwidth in 0.18 MHz ($15 \text{ kHz} \times 12$) PRB units. Constraints (11) arrange the lower and upper PRB limits of the numerology indexes included in profiles, while Constraints (12) avoid exceeding the overall bandwidth in every profile.

$$(\alpha_{p,g}^H - \alpha_{p,g}^L) \times E_g \geq M_{p,g}, \forall p \in \mathcal{P}, \forall g \in \bar{\mathcal{G}}_p \quad (13)$$

These constraints ensure a minimum bandwidth, $M_{p,g}$, required to ensure the packet delay budget for services with latency and GBR requirements (Assumption 1). The efficiency values, E_g , are given in Table 1.

Fairness inside profiles:

$$\Theta_p \leq (\alpha_{p,g}^H - \alpha_{p,g}^L) \times 2^g \times E_g \times \frac{1}{R_{p,g}}, \forall p \in \mathcal{P}, \forall g \in \bar{\mathcal{G}}_p \quad (14)$$

These constraints ensure that in each profile, the numerology indexes get fair BWPs. Note that the efficiency will be lower for numerology indexes with a large CP, so Assumption 3 is accounted for through the use of E_g , $0 \leq E_g \leq 1$. $R_{p,g}$ is a relative data rate required in numerology index g , under profile p , allowing for spectrum allocation to be proportional to service needs in \mathbf{r}^d (considering all DRBs better served by numerology g in use at profile p). This accounts for Assumption 2.

Non-negativity assignment to variables:

$$\alpha_{p,g}^H, \alpha_{p,g}^L, \Theta_p \in \mathbb{R}^+. \quad (15)$$

The CPLEX (IBM ILOG CPLEX Optimizer Version 12.8.) optimizer is used to optimally solve both steps of the numerology profile selection problem. Note that the first step of the numerology profile problem will be solved sporadically, when DRB requirements change significantly (which happens rarely, as stated in Section 3), and can always be solved offline. The second step would be solved when DRB data rates change significantly, but the developed mathematical formulation is a linear programming (LP) problem and can be solved in polynomial time, which means that it will be solved quickly even for large instances [32]. It can also be solved offline. In the other circumstances, the operators should switch between pre-planned profiles according to the currently present DRBs.

4.4. User BWP Assignment

Under each profile in operation, the user BWP allocation should be decided in real-time so that UE characteristics are taken into account, which can be ensured by the scheduler. Such allocation must not only take Assumptions 1–3 into account, but also the device energy constraints, as follows:

Assumption 4 (Energy). *For energy-constrained devices, the channel bandwidths will be low, and under normal conditions, the BWP/modulation should fit the expected throughput for energy efficiency. A short TTI (high numerology index) allows the device to enter sleep mode. When using high power (distance related) for transmission, the BWP/modulation should be small.*

5. Analysis of the Results

5.1. Relative Effect of Numerology Indexes

There are different supported scheduling approaches: slot based and non-slot based [33]. Moreover, the number of OFDM symbols per TTI can be different (Table 1). In other words, resources can be scheduled for users on a complete slot basis (14 OFDM symbols), or mini-slot scheduling can be used. A mini-slot is typically 2, 4, or 7 symbols [34]. A short TTI can be obtained by increasing the subcarrier spacing, Δf , or reducing the number of symbols per TTI. With this in mind and considering the information in Table 5, the relative effect of numerology indexes, e_g^d , is determined for the services in Table 4 as follows:

- VoIP, video conferencing, and real-time gaming: For these services, a GBR is required. Therefore, a short TTI duration is necessary. For video conferencing and real-time gaming, higher numerology indexes should be used due to their relatively high data rate requirements. Moreover, for real-time gaming, high numerologies can ensure the required low latencies. For VoIP, mini-slots in low numerologies can be used.
- Buffered streaming: In this type of application, the latency requirements are not strict. Therefore, a large TTI can be adopted. Whenever packet retransmission is needed, packet prioritization and/or buffering techniques can be used to cope with the required low packet error rate. Therefore, low numerologies should be used.
- Live streaming and pre-AR/VR: Since pre-AR/VR requires firm latency requirements and low packet error rates, Numerology Index 2 using extended CP can be used. Live streaming requires a smaller packet error rate, has no GBR requirements, and has some latency concerns. Therefore, middle numerology indexes (with normal CP) can be used.

The relative effect of numerology indexes for services in Table 4 is summarized in Table 6.

Table 6. Effect of numerology indexes for services in Table 4.

Service	$g = 0$	$g = 1$	$g = 2$	$g = 3$	$g = 4$
VoIP (if q is L/DR/GBR)	1 *	0.75	0.5	0.25	0
Video conference (if q is L/DR/GBR)	0	0.25	0.5	0.75	1
Real-time gaming (if q is L/DR/GBR)	0	0.25	0.5	0.75	1
Buffered streaming (if q is L/GBR)	1	0.75	0.5	0.25	0
Live streaming (if q is L/PER)	0	0.5	1	0.5	0
Pre-AR/VR (if q is L/PER)	0	0	1 **	0	0

L = latency; DR = data rate; GBR = guaranteed bit rate; PER = packet error rate. * Use of mini-slots. ** Use of extended cyclic prefix (CP).

5.2. Simulation Set Up

Let us assume an overall bandwidth $W = 50$ MHz, which is supported in multiple numerologies, corresponding to a total number of 270 PRBs in 0.18 MHz ($15 \text{ kHz} \times 12$) PRB units [23,35]. For simplicity, let us assume a fixed guard band of $B = 50$ kHz between any two numerologies, which is ≈ 0.28 in 0.18 MHz PRB units [12].

For services with latency and GBR requirements, the minimum bandwidth to ensure the packet delay budget, $M_{p,g}$, can be obtained as the following: Let $T_{\text{symbol}} = \frac{T_{\text{slot}}}{14}$ be the symbol duration, where T_{slot} is obtained from Table 1; the maximum number of allowed symbols in time, due to latency constraints, will be $\frac{L}{T_{\text{slot}}/14}$, and $M_{p,g}$ will be $\frac{GBR}{\frac{L}{T_{\text{slot}}/14} \times m}$ subcarriers, where m , GBR , and L are the maximum modulation factor, the total GBR needs of DRBs fitting numerology index g , and the average latency requirement of these DRBs, respectively. For the E_g values, these are shown in Table 1. Therefore:

$$M_{p,g} = \frac{GBR \times T_{\text{slot}}}{L \times 14 \times m \times 12}$$

in PRB units. Note that T_{slot} is dependent on the numerology. Regarding $R_{p,g}$, it can be obtained from the data rates shown in Table 4. The just mentioned simulation parameter values are summarized in Table 7.

Table 7. Simulation parameters.

Parameter	Value
Bandwidth, W	50 MHz
Guard band, B	50 kHz
Modulation order	1024
Available numerology indexes	0–4; see Table 1
DRB services	see Table 4
DRB's GBR and latency requirements	see Table 4

5.3. Discussion

5.3.1. Antagonistic DRBs and Numerologies

Tables 8 and 9 show the set of antagonistic DRBs and numerology indexes that better serve these antagonistic DRBs according to Table 6, for different profile granularities (maximum number of competing DRBs considered, k). The information in these tables basically includes the outcome of Step 1, from the two-step approach adopted to solve the numerology profile problem. Note that a high k value leads to fine-grained profiles.

According to these results, the most antagonistic DRBs are video conferencing and Pre-AR/VR, and then VoIP, buffered streaming, real-time gaming, and live streaming as the granularity of the profiles, k , increases. The numerology indexes that best meet the requirements of such applications are Numerology Index 2 with extended CP (2eCP) and 4 for $k = 2$. For $k = 3$, the best numerology indexes are 0, 2eCP, and 4. These three numerology indexes seem to be adequate for all applications except live streaming, and that is why no extra numerology index is selected until the live streaming is included as an antagonistic DRB, for which Numerology Index 2 is more adequate. Therefore, not all numerology indexes will be used, even if all applications are assigned to their best numerology index. Of course, this depends on the wireless communication scenario and kind of DRBs/applications under consideration. The numerology index that best meets the requirements of each application is shown in Table 10.

Table 8. Most antagonistic DRBs.

Application	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
VoIP	-	✓	✓	✓	✓
Video conference	✓	✓	✓	✓	✓
Real-time gaming	-	-	-	✓	✓
Buffered streaming	-	-	✓	✓	✓
Live streaming	-	-	-	-	✓
Pre-AR/VR	✓	✓	✓	✓	✓

Table 9. Selected numerologies.

Numerology Index	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
0	-	✓	✓	✓	✓
1	-	-	-	-	-
2	-	-	-	-	✓
2eCP *	✓	✓	✓	✓	✓
3	-	-	-	-	-
4	✓	✓	✓	✓	✓

2eCP * = Numerology Index 2 with extended CP.

Table 10. Numerology indexes that best meet the application requirements.

Application	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
VoIP	4	0	0	0	0
Video conference	4	4	4	4	4
Real-time gaming	4	4	4	4	4
Buffered streaming	4	0	0	0	0
Live streaming	4	0	0	0	2
Pre-AR/VR	2eCP	2eCP	2eCP	2eCP	2eCP

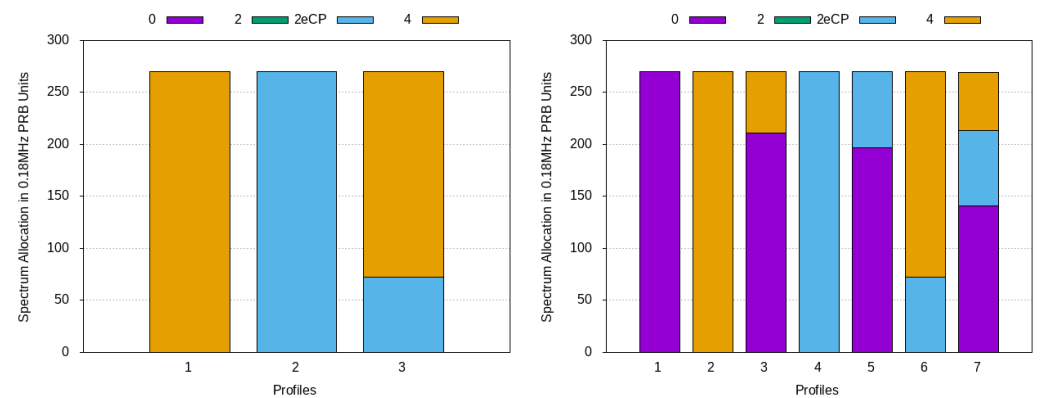
5.3.2. Spectrum Allocation

The plots in Figure 3 show the spectrum allocation in 0.18MHz ($15\text{kHz} \times 12$) PRB units. For $k = 4$ and $k = 5$, the set of numerology indexes in operation will be the same as for $k = 3$. In other words, although the set of possible profiles increases, some of them activate the same numerology indexes. For this reason, these are grouped when plotting spectrum allocation for $k = 6$.

Each of the $P = 2^k$ numerology profiles being planned can be represented by a k -bit pattern. An activated bit corresponds to the presence of an antagonistic DRB in the wireless communication scenario and is to be considered when distributing spectrum. Such a design of numerology profiles has the following advantages:

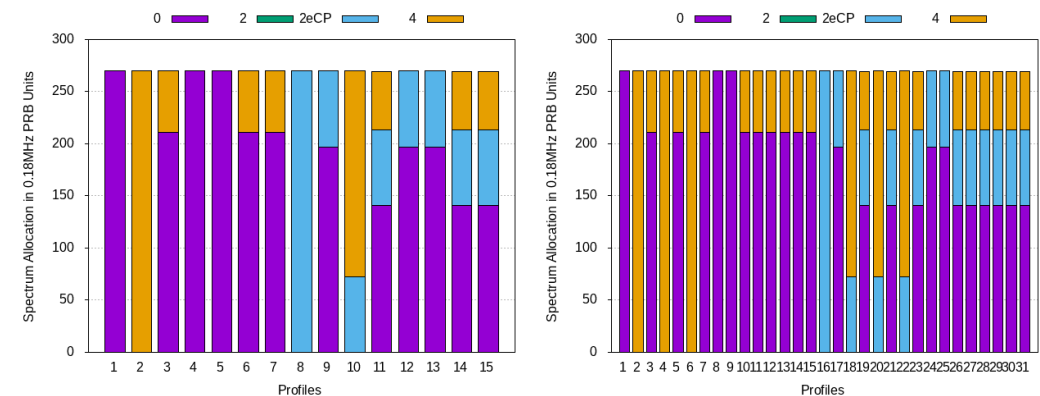
- Changes in users/services, currently present in a wireless communication scenario, can be followed by non-abrupt changes in spectrum allocation. That is, multiple transitions between numerology profiles, differing in a single bit, can be performed until the final numerology profile is reached. This allows for spectrum reconfiguration to be gradually introduced, avoiding posing problems to the users/services.
- When moving from one profile to another, at most one currently operating numerology index will have to be redefined to insert another numerology index.

These features allows for RRC reconfigurations (required for the new spectrum distribution to take place) to be kept minimized while ensuring full fairness of spectrum allocation among users/services.



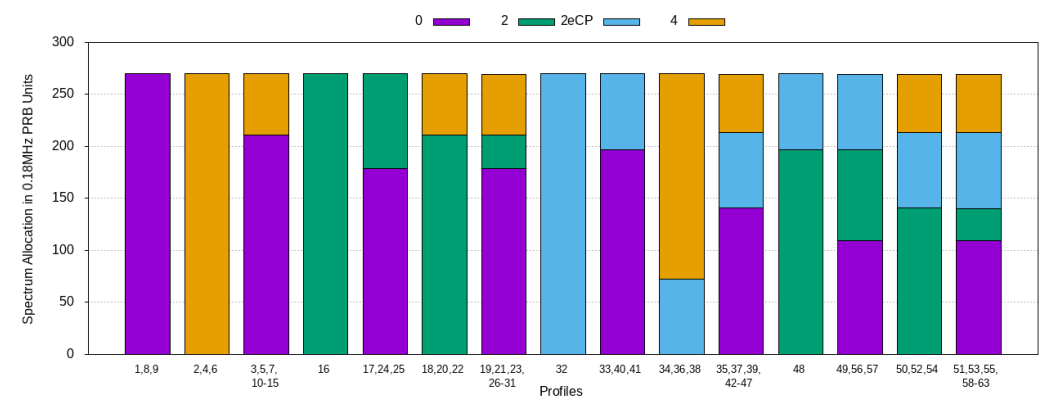
(a) Number of Antagonistic DRBs: $k = 2$

(b) Number of Antagonistic DRBs: $k = 3$



(c) Number of Antagonistic DRBs: $k = 4$

(d) Number of Antagonistic DRBs: $k = 5$



(e) Number of Antagonistic DRBs: $k = 6$

Figure 3. Spectrum allocation.

5.3.3. Overall Spectral Efficiency

The plots in Figure 4 show the spectral efficiency in the determined profiles. In general, the greater the amount of spectrum allocated to numerology index 2eCP, the lower the spectral efficiency is, as expected. Please note that these plots refer to the overall spectral efficiency, and fair spectrum allocation among the numerology indexes is guaranteed. This is ensured by the expression (14) included in the mathematical optimization solving the second step of the numerology profile problem. Fairness is proportional to the relative data rate need in operating numerology indexes. This ensures equal quality of experience for the different users/services.

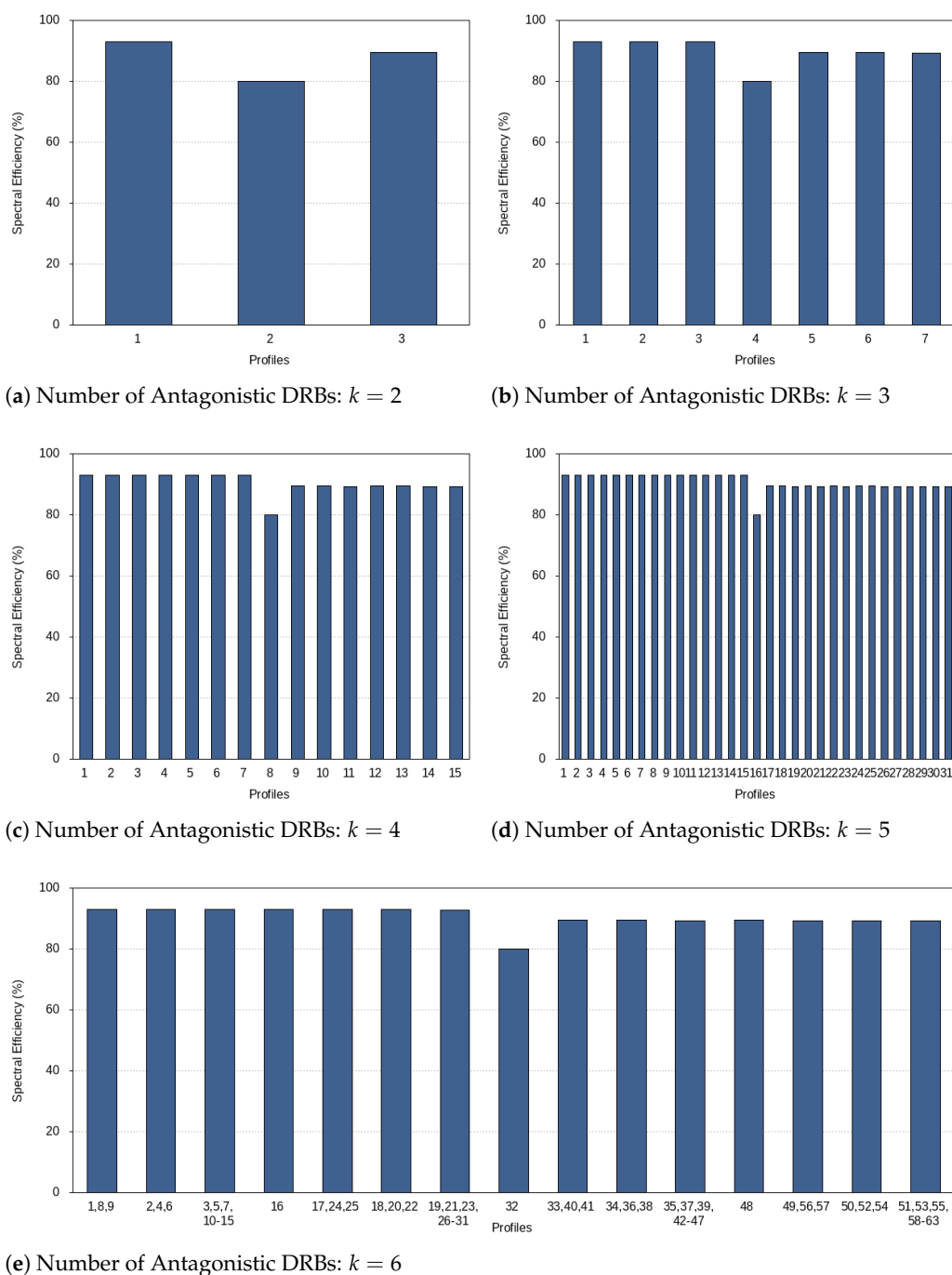


Figure 4. Overall spectral efficiency.

5.3.4. Satisfaction of QoS Requirements

The plots in Figure 5 show the average satisfaction of user/application requirements. The “optimized” bar is obtained by applying $\Phi(d, g, q) = r^d(q) - r^d(q) \times e_g^d(q)$ to the numerology index that best fits DRB d , from the numerology indexes available in a certain profile, and averaging the values obtained for all requirements, $q \in \{1, \dots, Q_d\}$. That is, according to the traffic types under consideration, shown in Table 4, q will be latency, average data rate (DR), GBR, and PER. The “non-optimized” bar is obtained by taking the least suited numerology index. Note that the average satisfaction goes from 0% to 100% because $0 \leq e_g^d(q), r^d(q) \leq 1$.

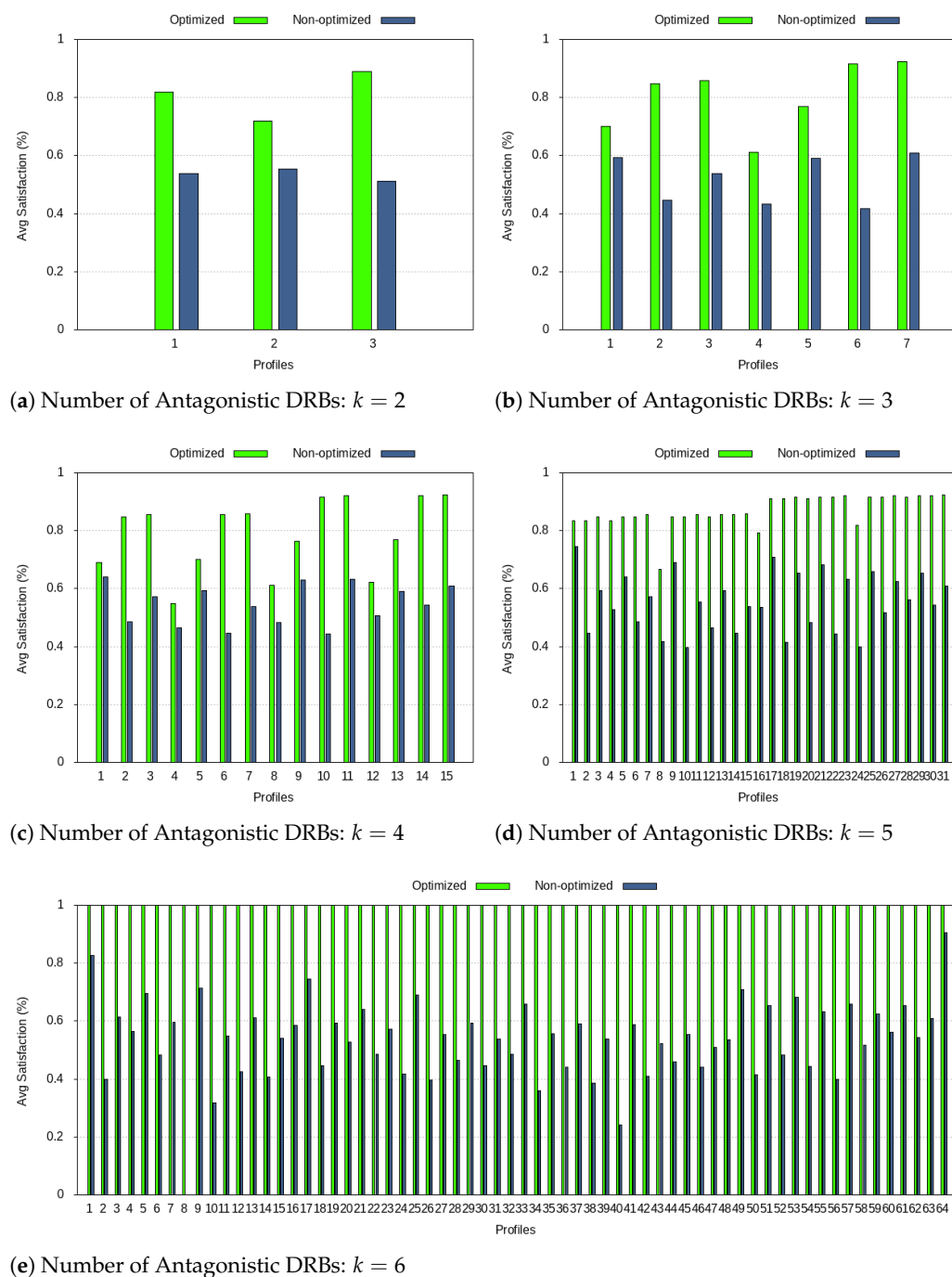


Figure 5. Average satisfaction of QoS requirements.

Results show that the optimization approach is able to choose adequate numerology indexes. The ranges for the “optimized” and “non-optimized” satisfaction percentage values are [0.6–1] and [0.4–0.65], respectively, and the values improve as the number of antagonistic DRBs considered, k , increases. Depending on the profile, the satisfaction level of the “optimized” approach can be twice as much as the “non-optimized” one. Therefore, the mathematical optimization is able to find numerology indexes that better meet user/service requirements, while keeping the total number of numerology indexes in operation minimized. For $k = 6$, one-hundred percent satisfaction is reached because the mixed numerology includes all required numerology indexes, fitting all requirements adequately.

5.4. Applicability of the Model

According to TS38.321, the RRC method is used for the configuration of a BWP set at any stage of the call, but the processing of RRC messages can reach 10 ms. On the other hand, the DCI-based adaptation allows on-the-fly activation/switching of a BWP, and the latency can be as low as 2 ms. Since RRC reconfigurations are the most demanding ones, the initially configured BWPs should fit not only current user/service needs (while being efficient regarding the use of physical resources), but also future needs, for an on-the-fly activation of BWPs. Besides BWPs, the network also informs a UE of the cell bandwidth.

The proposed framework is useful for the planning of BWPs on the network side. The use of profiles helps define adequate BWP sets, to be configured in a UE at the beginning of calls. This promotes on-the-fly activation of BWPs during the call, instead of RRC reconfigurations. The proposed framework can be applied to any wireless communication scenarios because the relative effect of each numerology index, on each QoS requirement, is set individually, and these can be adjusted according to the impairments of the scenario. Since the requirements of DRBs, typically present in the communication scenario, are not expected to change much, such planning of long-term profiles is feasible.

6. Conclusions and Future Work

This article addresses the mathematical optimization of numerology profiles. These numerology profiles reflect the anticipation of possible competitions for spectrum among the most antagonistic DRBs. The developed mathematical model is flexible, using all degrees of freedom to reach the best multiplexing, depending on the presence (or not) of the DRBs. The results show that the numerology indexes under utilization and spectrum allocation are fully optimized, according to users/service needs, while minimizing the number of reconfigurations required for the accommodation of new DRBs, avoiding disrupting other users/services. Furthermore, depending on the granularity of the profiles, the proposed optimization model is able to reach satisfaction levels of 60–100%, whereas a non-optimized approach reaches 40–65%, while simultaneously minimizing the total number of numerology indexes in operation. The future work will focus on developing a machine learning solution for profile selection and transition. Since profile selection can be converted into a sequential decision problem, we will develop an adaptive reinforcement learning solution to this problem and study it under different use cases and deployment scenarios.

Author Contributions: Conceptualization, methodology, formal analysis, software, validation, and writing: N.C. and F.A.-T.; investigation: N.C., F.A.-T., and J.R.; supervision and funding acquisition: N.C. and J.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the European Regional Development Fund (FEDER), through the Competitiveness and Internationalization Operational Programme (COMPETE 2020), the Regional Operational Program of the Algarve (2020), and Fundação para a Ciência e Tecnologia; i-Five: Extensão do acesso de espectro dinâmico para rádio 5G, POCI-01-0145-FEDER-030500. This work is also supported by Fundação para a ciência e Tecnologia within CEOT (Center for Electronic, Optoelectronic and Telecommunications) and the UID/MULTI/00631/2020 project.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- David, K.; Berndt, H. 6G Vision and Requirements: Is There Any Need for Beyond 5G? *IEEE Veh. Technol. Mag.* **2018**, *13*, 72–80. [CrossRef]
- Navarro-Ortiz, J.; Romero-Diaz, P.; Sendra, S.; Ameigeiras, P.; Ramos-Munoz, J.J.; Lopez-Soler, J.M. A survey on 5G usage scenarios and traffic models. *IEEE Commun. Surv. Tutor.* **2020**, *2*, 905–929. [CrossRef]
- 3GPP. TSG RAN; NR; Overall Description; Stage 2 (Release 15), v15.3.0, 3GPP TS 38.300; Technical Report; ETSI: Sophia Antipolis, France, 2018. Available online: <https://www.3gpp.org/release-15> (accessed on 21 February 2021).
- Jeon, J. NR Wide Bandwidth Operations. *IEEE Commun. Mag.* **2018**, *3*, 42–46. [CrossRef]
- 3GPP. Specification # 38.201: NR; Physical Layer; General Description (Release 15); Technical Report; ETSI: Sophia Antipolis, France, 2017. Available online: <https://www.3gpp.org/release-15> (accessed on 21 February 2021).
- Zaidi, A.; Baldemair, R.; Tullberg, H.; Björkegren, H.; Sundstrom, L.; Medbo, J.; Kilinc, C.; Da Silva, I. Waveform and Numerology to Support 5G Services and Requirements. *IEEE Commun. Mag.* **2016**, *54*, 90–98. [CrossRef]
- Marijanovic, L.; Schwarz, S.; Rupp, M. Optimal Resource Allocation with Flexible Numerology. In Proceedings of the IEEE International Conference on Communication Systems (ICCS), Chengdu, China, 19–21 December 2018, pp. 136–141. [CrossRef]
- Parkvall, S.; Dahlman, E.; Furuskar, A.; Frenne, M. NR: The New 5G Radio Access Technology. *IEEE Commun. Stand. Mag.* **2017**, *1*, 24–30. [CrossRef]
- Zhang, L.; Ijaz, A.; Xiao, P.; Quddus, A.; Tafazolli, R. Subband Filtered Multi-Carrier Systems for Multi-Service Wireless Communications. *IEEE Trans. Wirel. Commun.* **2017**, *16*, 1893–1907. [CrossRef]
- Kihero, A.; Solaija, M.J.; Yazar, A.; Arslan, H. Inter-Numerology Interference Analysis for 5G and Beyond. In Proceedings of the 2018 IEEE Globecom Workshops (GC Wkshps), Abu Dhabi, United Arab Emirates, 9–13 December 2018. [CrossRef]
- Bag, T.; Garg, S.; Shaik, Z.; Mitschele-Thiel, A. Multi-Numerology Based Resource Allocation for Reducing Average Scheduling Latencies for 5G NR Wireless Networks. In Proceedings of the 2019 European Conference on Networks and Communications (EUCNC), Abu Dhabi, United Arab Emirates, 9–13 December 2019; pp. 597–602.
- Yazar, A.; Arslan, H. A Flexibility Metric and Optimization Methods for Mixed Numerologies in 5G and Beyond. *IEEE Access* **2018**, *6*, 3755–3764. [CrossRef]
- You, L.; Liao, Q.; Pappas, N.; Yuan, D. Resource Optimization With Flexible Numerology and Frame Structure for Heterogeneous Services. *IEEE Commun. Lett.* **2018**, *22*, 2579–2582. [CrossRef]
- Marijanovic, L.; Schwarz, S.; Rupp, M. A Novel Optimization Method for Resource Allocation based on Mixed Numerology. In Proceedings of the ICC 2019–2019 IEEE International Conference on Communications (ICC), Shanghai, China, 20–24 May 2019.
- Nguyen, T.T.; Ha, V.N.; Le, L.B. Wireless Scheduling for Heterogeneous Services With Mixed Numerology in 5G Wireless Networks. *IEEE Commun. Lett.* **2020**, *24*, 410–413. [CrossRef]
- Tang, J.; Shojaeifard, A.; So, D.K.C.; Wong, K.K.; Zhao, N. Energy Efficiency Optimization for CoMP-SWIPT Heterogeneous Networks. *IEEE Trans. Commun.* **2018**, *66*, 6368–6383. [CrossRef]
- Anand, A.; De Veciana, G.; Shakkottai, S. Joint scheduling of URLLC and eMBB traffic in 5G wireless networks. *IEEE/ACM Trans. Netw.* **2020**, *28*, 477–490. [CrossRef]
- Xu, L.; Yu, X.; Gulliver, T. Intelligent Outage Probability Prediction for Mobile IoT Networks Based on an IGWO-Elman Neural Network. *IEEE Trans. Veh. Technol.* **2021**. [CrossRef]
- Zhang, J.; Xu, X.; Zhang, K.; Zhang, B.; Tao, X.; Zhang, P. Machine Learning Based Flexible Transmission Time Interval Scheduling for eMBB and uRLLC Coexistence Scenario. *IEEE Access* **2019**, *7*, 65811–65820. [CrossRef]
- Correia, N.; Al-Tam, F.; Rodriguez, J. Adaptive Spectrum Allocation for 5G Wireless Communication Scenarios. In Proceedings of the 2020 IEEE 25th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD), Pisa, Italy, 14–16 September 2020.
- Zhang, X.; Zhang, L.; Xiao, P.; Ma, D.; Wei, J.; Xin, Y. Mixed Numerologies Interference Analysis and Inter-Numerology Interference Cancellation for Windowed OFDM Systems. *IEEE Trans. Veh. Technol.* **2018**, *67*, 7047–7061. [CrossRef]
- Huang, Y.; Li, S.; Li, C.; Hou, Y.T.; Lou, W. A Deep-Reinforcement-Learning-Based Approach to Dynamic eMBB/URLLC Multiplexing in 5G NR. *IEEE Internet Things J.* **2020**, *7*, 6439–6456. [CrossRef]
- 3GPP. TS 38.101-1-g10: User Equipment (UE) Radio Transmission and Reception; Part 1: Range 1 Standalone (Release 15); Technical Report; ETSI: Sophia Antipolis, France, 2018. Available online: <https://www.3gpp.org/release-15> (accessed on 21 February 2021).
- ITU. IMT Vision—Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond (Recommendation ITU-R M.2083-0); Technical Report; ITU-R: Geneva, Switzerland, 2015. Available online: <https://www.itu.int/rec/R-REC-M.2083-0-201509-I> (accessed on 21 February 2021).
- Yu, H.; Leeand, H.; Jeon, H. What is 5G? Emerging 5G Mobile Services and Network Requirements. *Sustainability* **2017**, *10*, 1848.

-
26. Zikria, Y.; Kim, S.; Afzal, M.; Wang, H.; Rehmani, M. 5G Mobile Services and Scenarios: Challenges and Solutions. *Sustainability* **2018**, *10*, 3626. [[CrossRef](#)]
 27. Tseng, S.C.; Liu, Z.W.; Chou, Y.C.; Huang, C.W. Radio Resource Scheduling for 5G NR via Deep Deterministic Policy Gradient. In Proceedings of the IEEE International Conference on Communications Workshops (ICC Workshops), Shanghai, China, 20–24 May 2019.
 28. 3GPP. TS 23.501: *System Architecture for the 5G System*; Technical Report; ETSI: Sophia Antipolis, France, 2018. Available online: <https://www.3gpp.org/release-15> (accessed on 21 February 2021).
 29. 3GPP. Specification #3GPP TR 38.913: *5G; Study on Scenarios and Requirements for Next Generation Access Technologies*; Technical Report; ETSI: Sophia Antipolis, France, 2017. Available online: <https://www.3gpp.org/release-15> (accessed on 21 February 2021).
 30. Pedersen, K.I.; Berardinelli, G.; Frederiksen, F.; Mogensen, P.; Szufarska, A. A Flexible 5G Frame Structure Design for Frequency-Division Duplex Cases. *IEEE Commun. Mag.* **2016**, *54*, 53–59. [[CrossRef](#)]
 31. Pedersen, K.; Poczni, G.; Steiner, J.; Maeder, A. Agile 5G Scheduler for Improved E2E Performance and Flexibility for Different Network Implementations. *IEEE Commun. Mag.* **2018**, *56*, 210–217.
 32. Spielman, D.; Teng, S.H. Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time. *J. ACM* **2004**, *51*, 385–463. [[CrossRef](#)]
 33. Chandramouli, D.; Liebhart, R.; Pirskanen, J. *5G for the Connected World*; Wiley Online Library: Hoboken, NJ, USA, 2019.
 34. Harri Holma, A.T.; Nakamura, T. *5G Technology: 3GPP New Radio*; Wiley: Hoboken, NJ, USA, 2019.
 35. 3GPP. 5G; NR; *User Equipment (UE) Radio Transmission and Reception; Part 2: Range 2 Standalone (3GPP TS 38.101-2 Version 15.3.0 Release 15)*; Technical Report; ETSI: Sophia Antipolis, France, 2018. Available online: <https://www.3gpp.org/release-15> (accessed on 21 February 2021).